**[네트워크 포렌식을 위한**

**패킷분석 및 데이터 복원]**

**(제안서)**

2016년 03월

호서대학교 공과대학 정보보호학과

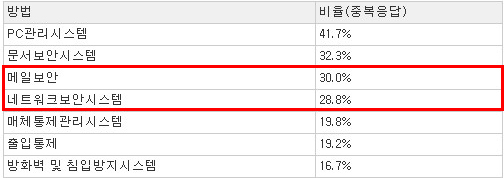
|  |  |  |
| --- | --- | --- |
| **제 출 문** | | |
| **과 제 명** | 네트워크 포렌식을 위한 패킷분석 및 데이터 복원 | |
| **참여학생** | 학 번 | 성 명 |
| 20111748 | 위푸름 |
| 20123381 | 유진영 |
| 20122343 | 김수진 |
|  |  |
| **수행기간** | 2016.03.01 ~ 2016.00.00 | |
| 본 보고서를 호서대학교 정보보호학과  정보보호 프로젝트 제안서로 제출합니다.  2016년 3월 14일  **호서대학교 정보보호학과 학과장 귀하** | | |

프로젝트의 필요성

최근 국내외에서 기업내부의 기밀정보를 유출하는 경우가 빈번하게 발생하고 있는데, 유출사항으로는 아래의 [그림1]과 같이 데이터 유출 다음으로 이메일 및 메신저를 통한 유출이 많음을 확인할 수 있다. 이러한 유출은 기업의 막대한 손실을 불러오며 나아가 핵심기술의 경우에는 국가적 손실까지 불러올 수 있는 위험성이 존재한다.

[그림 1] 기밀정보 유출 경로

하지만 현실은 기업의 담당자는 물론 일반 직원들은 보안에 대한 인식이 저조할 뿐 아니라 중소기업의 경우 교육이 기업내의 보안교육이 제대로 이루어지지 않은 경우가 대부분이며 이에 지속적으로 문제점이 발생하였다. 기업들의 유출방지 대응방안으로는 아래 [그림2]와 같이 가장 많이 적용된 PC관리시스템도 50%가 넘지 않는 상태이며, 그 중에서도 메일 보안 및 네트워크보안시스템의 경우는 겨우 30%정도의 저조한 상태를 보여주고 있다.

[그림2] 기술유출방지 대응 시스템

그나마도 기술유출방지를 위해 인프라가 구축이 되어있고, 각종 보안프로그램이 설치되어 있다 해도 많은 사내 규정과 승인절차에 사용자가 불편함을 느껴, 설치하였다가 삭제하거나, 설치하지 않는 경우도 없지 않다. 또한 기업의 기밀을 유출시키는 원인인 경쟁사의 산업스파이의 경우 이러한 환경에서 사내 기밀문서를 유출할 확률이 높다.

데이터 유출 문제들을 해결하기 위해 디지털상의 자료(메시지 내용, 이메일, 로그 등)들도 법정에서 증거로 효력이 입증되면서부터 국가적인 차원에서 해당 분야에 대한 관심이 증가 하고 있다. 하지만 디지털 정보는 원본과 복사본의 구분이 어렵고, 위·변조가 용이하여 증거로서의 진정성을 의심하지 않을 수 없다. 따라서 디지털 정보가 증거로서의 가치를 인정받기 위해서는 정보의 수집·보존·처리 전 과정에서 특별한 절차와 방법을 필요로 한다. 따라서 디지털 정보를 근거로 하여 범죄행위의 사실관계를 규명하고 인과관계를 입증하는 과학적 수사기법인 디지털 포렌식이 요구된다

디지털 포렌식 혹은 컴퓨터 포렌식(Computer Forensic)은 전자적 증거물 등을 사법기관에 제출하기 위해 데이터를 수집, 분석, 보고서를 작성하는 일련의 작업을 말한다. 과거에 얻을 수 없었던 증거나 단서들을 제공해 준다는 점에서 획기적인 방법이다. 컴퓨터 포렌식은 사이버 해킹 공격, 사이버 범죄 시 범죄자들은 컴퓨터, 이메일, IT기기, 스마트폰 등의 운영체제, 어플리케이션, 메모리 등에 다양한 전자적 증거를 남기게 되면서, 사이버 범죄자 추적 및 조사에 핵심적인 요소가 되고 있다.

[표 1] 디지털 포렌식(Digital Forensics)유형

|  |  |
| --- | --- |
| 디지털 포렌식 유형 | 설 명 |
| 네트워크 포렌식  (Network Forensics) | 네트워크로 전송되는 데이터를 대상으로 증거 획득 및 분석 |
| 활성데이터 포렌식  (Volatile data Forensics) | 휘발성 데이터를 대상으로 증거 획득 및 분석 |
| 디스크 포렌식  (Disk Forensics) | 비 휘발성 저장매체(하드디스크, SSD, USB, CD 등)를 대상으로 증거 획득 및 분석 |
| 이메일 포렌식  (Email Forensics) | 이메일 데이터로부터 송/수신자, 송/수신 시간, 내용 등의 증거 획득 및 분석 |
| 웹 포렌식  (Web Forensics) | 웹 브라우저를 통한 쿠키, 히스토리, 임시파일, 설정 정보 등을 통해 사용 흔적 분석 |
| 모바일/임베디드 포렌식  (Mobile/Embedded Forensics) | 휴대폰, 스마트폰, PDA, 네비게이션, 라우터 등의 모바일 기기를 대상으로 증거 획득 및 분석 |
| 멀티미디어 포렌식  (Multimedia Forensics) | 디지털 비디오, 오디오, 이미지 등의 멀티미디어 데이터에서 증거 획득 및 분석 |
| 소스코드 포렌식  (Source code Forensics) | 프로그램 실행 코드와 소스 코드의 상관관계 분석, 악성코드 분석 |
| 데이터베이스 포렌식  (Database Forensics) | 방대한 데이터베이스로부터 유효한 증거 획득 및 분석 |
| 안티포렌식  (Anti-Forensics) | 데이터 완전 삭제, 암호화, 심층암호(Steganography) |

위의 [표 1]과 같이 디지털 포렌식의 유형 중 네트워크 포렌식은 패킷에 대한 심층분석을 해야 하는 만큼 중요한 기술 임에도 불구하고 국내에서는 아직까지 많은 연구가 이루어지지 않았으며, 더욱이 해당 분야에 대한 인력도 충분하지 않은 상태이며 그것을 활용한 실무 분야는 조금씩 활성화 되어가고 있지만, 아직까지 기업에서 그것을 활용한 감사는 진행되지 않고 있다. 따라서 기업의 유출된 문서에 대해서는 리스트를 줄이기 위하여 사후 분석 과정이 필요하고 해당 공격패턴분석을 통해 시그니처를 정립해 추후에 공격을 막을 수 있다.

프로젝트 목표 및 내용

2.1 프로젝트 목표

국내외에서 기업 내부의 기밀정보가 이메일·메신저 등 네트워크를 통해 유출되는 경우가 빈번하게 발생하고 있지만 네트워크를 통한 기술유출 대응방안 연구는 미흡한 실정이다. 따라서 본 프로젝트의 결과물인 파일 복원기능을 통해 기업의 보안 담당자 혹은 감사원이 유출된 파일에 대한 리스크 분석을 용이하게 하고, 국가별 IP traffic 통계 자료를 통해 기술유출 경로 파악을 용이하게 한다. 또한 I/O Traffic 통계 및 최다 사용 프로토콜 정보를 제공함으로써 기업 내 네트워크 품질 개선을 용이하게 하고자 한다.

2.2 프로젝트 내용

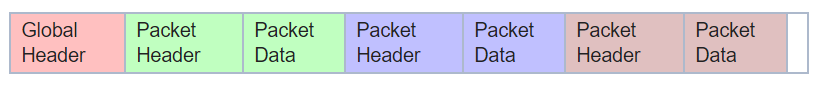
2.2.1 서비스 방식

본 프로그램은 Cross Platform을 지원하고, Web 어플리케이션의 형태로 제공된다. 이 때, 분석대상이 되는 덤프파일 사이즈가 클 경우 사용자의 프로그램 접근성이 떨어지게 된다. 이에 따라 제품을 패키지화하여 사용자가 자신의 PC에 패키지를 설치하고 localhost의 지정된 포트를 통해 설치된 패키지를 이용할 수 있게 함으로써 사용자가 패키지를 자신의 개인PC나 서버에 선택적으로 설치할 수 있게 한다. 이는 사용자가 프로그램을 자신의 스타일에 따라 설치하여 사용할 수 있다는 것을 의미한다.

2.2.2 전체 프로세스

본 프로그램의 전체적인 프로세스는 다음과 같다.

  
**[그림 2-1] 프로그램의 전체 흐름도**

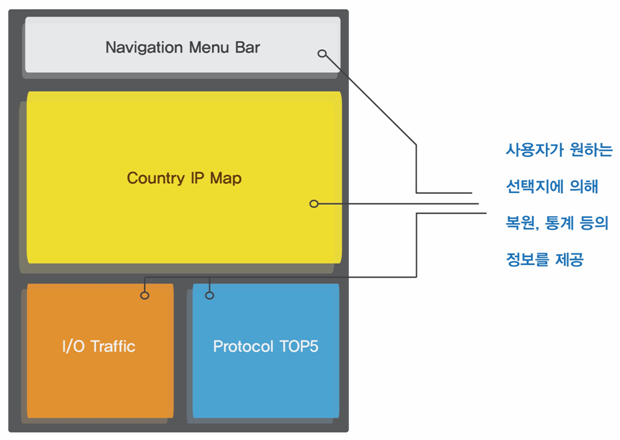
우선 사용자가 패킷 덤프파일을 업로드 하게 되면 대시보드의 그래프를 출력하기 위해 업로드 된 덤프파일에서 관련 정보를 추출해야 한다. 이를 위해 우선 덤프파일의 파싱 작업이 수행되는데 이는 업로드 된 파일이 분석 가능한 덤프파일인지 판별하고, 대표적인 덤프파일의 포맷인 pcap과 pcap-ng포맷을 구분한 후에 각 패킷을 분할하여 파싱 작업을 수행한다. 다음은 PCAP 파일의 구조를 그림으로 나타낸 것이다.  


**[그림 2-2] PCAP File Structure**

우선 Global Header 부분에는 PCAP파일인지 PCAP-NG파일인지를 나타내는 Magic Number가 포함 되어있다. PCAP파일의 Magic Number는 0xA1B2C3D4이고 PCAP-NG파일은 0x1A2B3C4D이다. 따라서 이것을 확인하면 정상적인 덤프파일인지 여부를 확인 할 수 있으며 추가로 PCAP과 PCAP-NG 포맷을 구분할 수 있다.  
바로 뒤부터는 패킷들의 헤더와 해당 패킷의 데이터가 존재한다. 이를 순차적으로 프로그램에서 Loading 하여 패킷들의 정보를 읽어와 분석 루틴을 수행할 수 있는데 이 때 분석루틴은 각 패킷이 모두 동일하지 않으므로 개별적으로 분석하여야 한다. 예를 들어 각 패킷별로 Protocol 정보가 다를 수 있으므로 각 Protocol별로 파싱 모듈을 제작하여 Protocol 구분 후에 분기하여 각 모듈별로 별도 분석을 하여야 한다.  
이러한 루틴으로 분석을 한 이후에는 대시보드 뿐만 아니라 그래프로 나타내야 할 정보들을 추출할 수 있다. 이렇게 추출된 정보를 분석결과라고 한다.  
분석결과를 도출하였다면 화면에 표시할 그래프별로 구조화 작업을 수행한다. 구조화 작업이란 Back-End 단에서 도출된 분석결과를 Front-End 단으로 전송하기 위함과 동시에 해당 분석결과를 사용자가 Save 및 Load 할 수 있도록 기능을 제공하기 위한 작업을 말한다. 분석결과를 구조화 할 때는 XML파일형태로 파일을 생성한다. XML파일로 분석결과를 구조화하면, Front-End와의 연동작업 시 간편하게 많은 정보를 체계적으로 전송할 수 있으며, Front-End 단에서 사용자의 분석파일을 Load하여 보여줄 때에도 통일된 모듈로 개발이 가능하다.

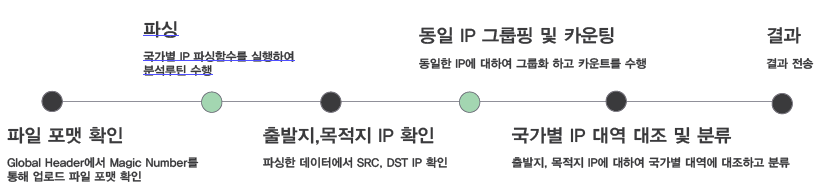
2.2.3 대시보드 화면

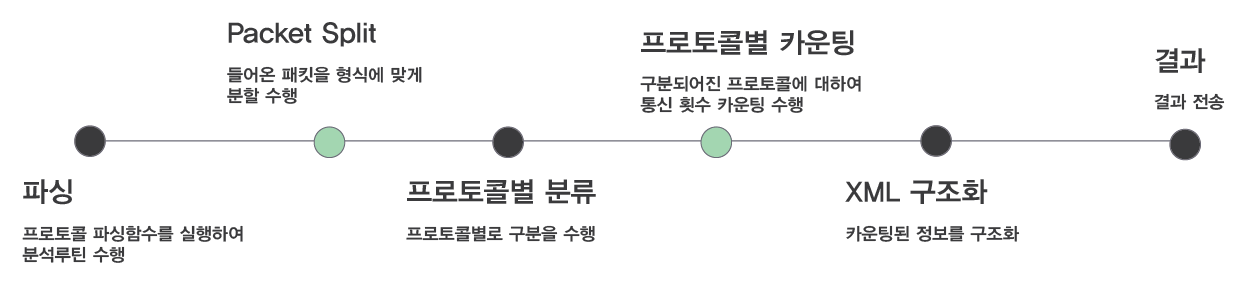
분석 후 사용자에게 제일 처음으로 보여지는 화면이며, 다음 그림과 같이 구성되어 있다.



**[그림 2-3] 대시보드 UI 구성**

맨 상단에는 네비게이션 바가 위치하고 있다. 이곳에서 사용자가 원하는 형태의 통계정보를 선택적으로 제공받을 수 있다.  
다음으로 국가별 IP 대역 분류 정보가 제공된다. 이곳에서는 세계지도가 렌더링되며 트래픽이 집중된 국가일수록 진한 색으로 표기가 된다. 추가적으로 Inbound Traffic, Outbound Traffic을 구분하여 보여주어 In/Out Traffic의 흐름을 파악 할 수 있다. 국가별 IP 대역 분류는 다음과 같은 프로세스로 진행된다.

  
**[그림 2-4] 국가별 IP 대역 분류 프로세스**

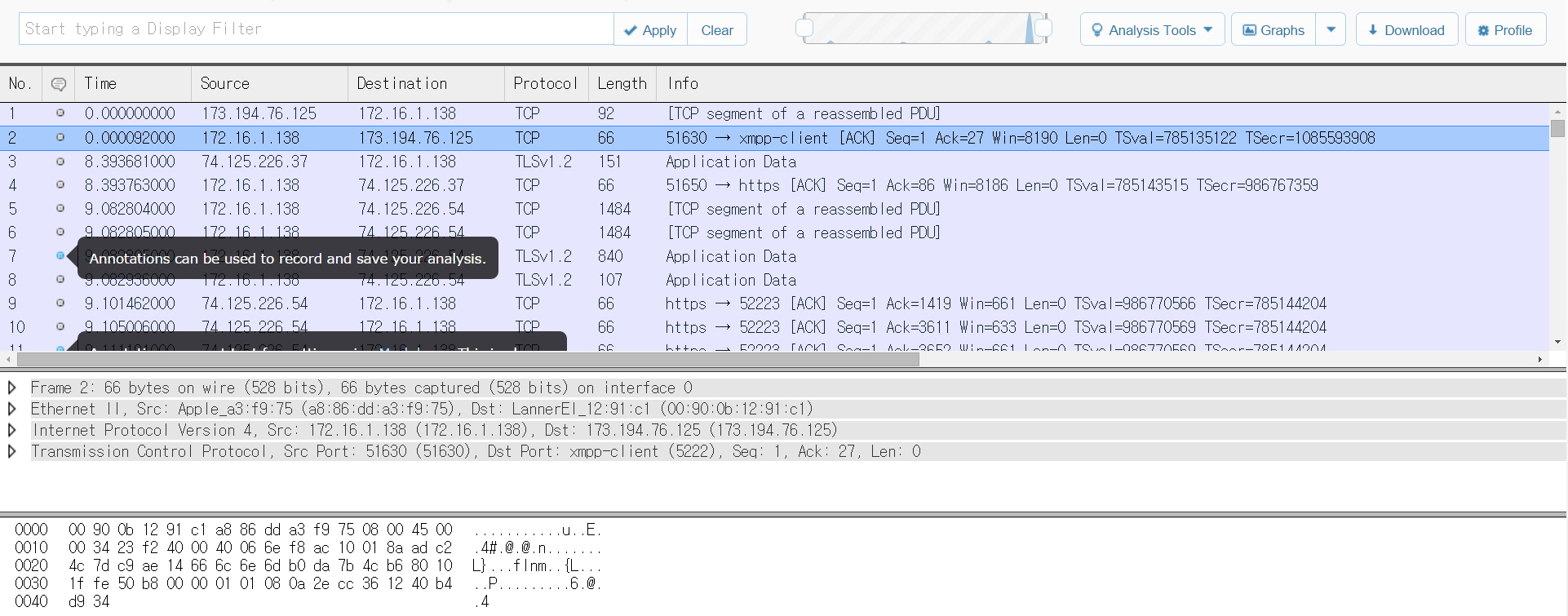
국가별 IP 대역 분류의 하단에는 Input 및 Output 트래픽의 비율을 원형그래프의 형태로 보여준다. 사용자는 이를 이용해 덤프파일의 패킷흐름을 대략적으로 예측할 수 있다.  
다음으로 우측에는 덤프파일 내 패킷들의 프로토콜 형태를 분류하여 가장 많이 사용된 5가지의 프로토콜을 보여준다. 

**[그림 2-5] top5 프로토콜 분석 프로세스**

사용자는 최종적으로 대시보드 화면을 통해 In/Out Traffic과 사용 Protocol에 대한 흐름 파악이 가능하다.

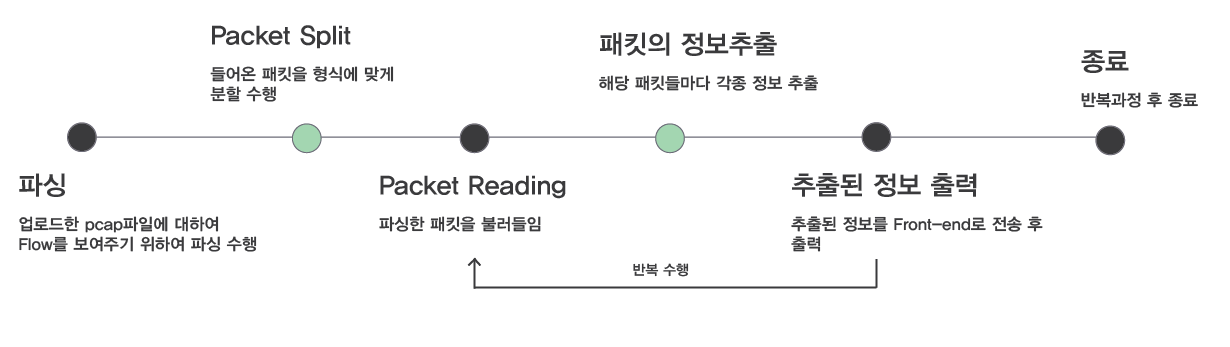
2.2.4 Flow 화면

추가적으로 사용자에게 제공되는 화면은 Flow화면이다. Flow화면은 분석을 위해 패킷별로 보다 자세한 정보를 제공하기 위한 화면이다. 이는 기존의 제품들에서 찾아볼 수 있는 화면과 동일하며 예시화면은 다음과 같다.



**[그림 2-6] CloudShark의 패킷 캡쳐 화면**

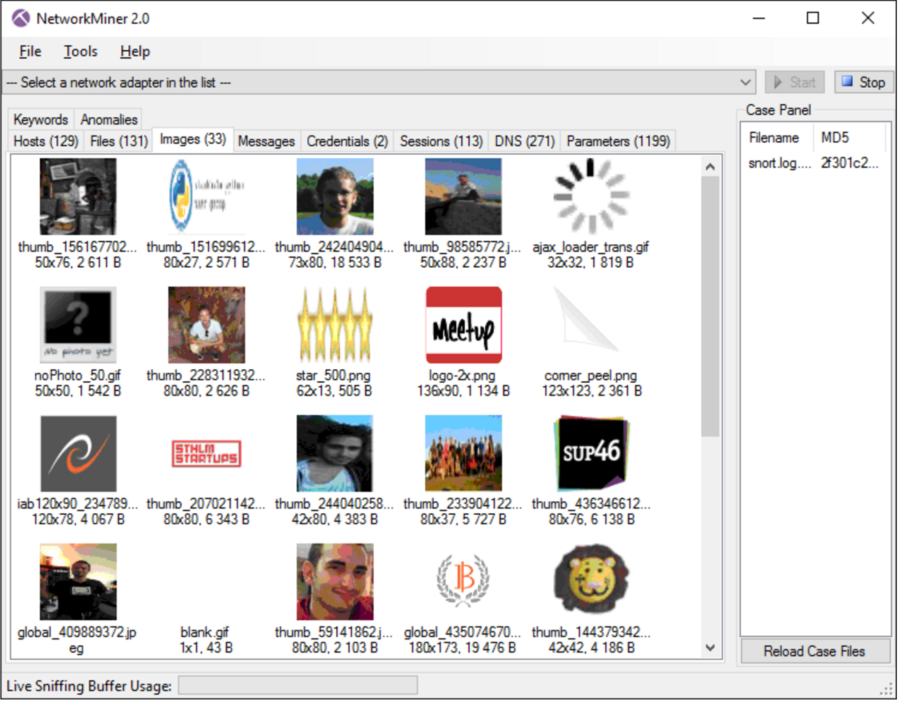
Web Application으로 서비스되는 CloudShark의 패킷 분석 화면이다. 패킷의 정보를 테이블형식으로 Time, Source IP, Destination IP, Protocol, Length, Info의 필드로 나누어 분류하여 제공한다. 본 제품도 마찬가지로 테이블형식으로 제공하며 제공되지만 기존 제품보다 더욱 자세한 정보가 제공된다. 예를 들어, Source IP, Destination IP, Domain Name, Source Port, Destination Port, TCP/UDP구분, Protocol 종류, Traffic 발생 국가, 패킷의 사이즈, 패킷의 길이 등이 제공될 예정이다.  
 본 Flow화면을 구현할 때는 순차적으로 패킷을 분할하여 분석하여 화면에 나타낸다. 왜냐하면 덤프파일의 사이즈가 클 경우 분석시간이 오래 걸릴 수 있으므로 사용자 입장에서 기다리는 시간을 줄일 수 있게 된다.

.**[그림 2-7] Flow 분석 프로세스**

.

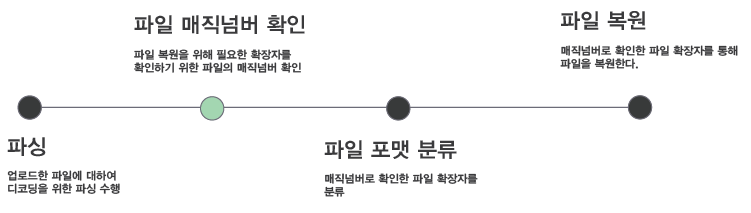
2.2.5 파일복원

본 프로그램에서는 분석기능에 추가적으로 파일 복원기능을 제공한다. 다음은 파일 복원이 지원되는 패킷 분석 도구인 NetworkMiner로 이미지 복원을 수행한 화면이다.



**[그림 2-8] NetworkMiner의 이미지 복원**

NetworkMiner는 보이는 것처럼 이미지와 파일을 따로 탭으로 구분하여 준다. 복원된 파일이 이미지일 경우 썸네일 형식으로 사용자가 다운로드 해볼 필요 없이 미리 확인 해 볼 수 있다. 따라서 본 제품도 파일을 분류하여 사용자가 확인하기 편리하도록 하며, 추가적으로 이미지 썸네일 기능을 제공한다.  
분석기능에서의 프로세스는 다음과 같다.



**[그림 2-9] 파일 복원 프로세스**

우선 덤프파일을 2.2.2 절에서 설명한 것과 같이 파싱 작업을 수행한다. 패킷별로 파싱 작업을 수행한 이후에는 파일의 매직넘버를 확인한다. 파일은 확장자 별로 고유의 매직넘버를 가진다. 따라서 매직넘버를 확인하여 대조하면 파일의 확장자를 알 수 있으므로 분류가 가능하다. 이 때 분류된 파일 중 이미지파일 확장자만 따로 분류하여 썸네일 형식으로 사용자들에게 보여지게 된다.  
파일을 복원 할 때는 패킷에서 파일의 바이너리코드(Binary Code)를 Loading 하여 별도의 파일에 해당 바이너리 코드를 쓴다. 그 이후 매직넘버에 의해 판별된 확장자를 추가하여 저장 해주면 파일 복원이 완료된다.  
파일 복원을 수행할 때는 Flow화면을 구현할 때와 마찬가지로 Global Header 다음의 패킷부터 순차적으로 분석 루틴을 수행하고 파일을 복원하는 즉시 화면에 보여준다. 왜냐하면 Flow화면을 구현할 때와 동일한 이유로 덤프파일의 사이즈가 큰 경우에 분석 속도가 오래 걸리는 문제를 해결하기 위한 방법이다.

프로젝트의 추진전략 및 방법

3.1 개발 환경

OS: window 10

사용 언어: python, c, css, javascript, html

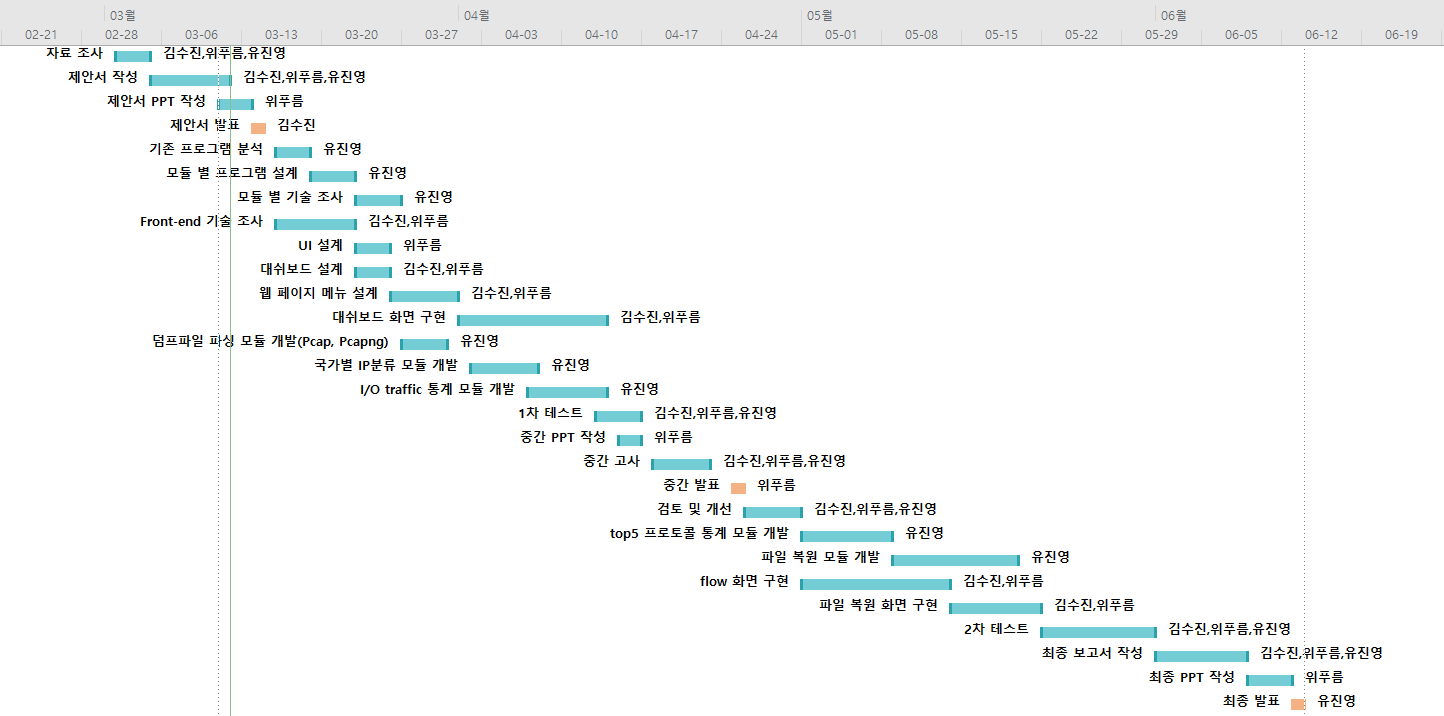
사용 프로그램: VMware, github, sourceinsight, sublime text

3.2 역할 분담

|  |  |
| --- | --- |
| 위푸름 | UI·UX 디자인 및 PPT제작, 대시보드 화면 구현,세부 메뉴 화면 구현 |
| 유진영 | 파싱 모듈 개발, 통계 모듈 개발, 복원 모듈 개발 |
| 김수진 | 대시보드 화면 구현, 세부 메뉴 화면 구현 |

3.3 일정

3.3.1 향후 계획



3.3.1 상세 일정



프로젝트 결과의 활용 방안 및 기대효과

기업의 보안담당자 혹은 감사원은 의심되는 이용자에 대한 색출과 이용자가 한 행위(파일 전송, 메일전송 등)를 파악하기 위하여 기업 내에 있는 보안관제부서에 해당 이용자의 행위에 대한 네트워크 패킷을 요구하여 분석을 실시할 수 있다. 회사 내규 등에 임직원 프라이버시 침해가 아닌 기업 리스크 관리를 위한 모니터링의 취지를 밝히고 임직원 동의서를 받은 후 네트워크 패킷 분석을 통한 모니터링으로 접근 사이트 주소, 이메일 수신 및 발신 기록 등 다양한 분야에서 내부통제 강화 용도로 활용 된다. 또한 다양한 통계 데이터(I/O traffic 통계, 최다 사용 프로토콜 통계 등)를 이용해 기업 내 네트워크 품질 개선 자료로 활용될 수 있다.

네트워크 패킷 분석을 통한 데이터 복원 기술은 패킷의 파일 및 프로토콜 정보를 이용하여 원래의 파일의 형태로 복원하기 위한 기술이다. 이 기술을 이용하여 유출된 파일에 대한 리스크 분석과 사후 처리 용도로 활용된다. 또한 파일이 아닌 다양한 제어 패킷들도 수집/복원되고, 이런 정보들을 통해서 비인가 사설 FTP서버 접속 탐지, 특정 포트를 이용한 비인가 서 비스 이용 탐지 등과 같은 보안사고를 예방할 수 있다.

IPS와 방화벽의 경우 보통 시그니처 기반으로 탐지하여 공격의 유무 판단을 위주로 수행을 한다면 본 프로젝트의 프로그램은 그 행위에 대하여 더 나아가 패킷 심층분석을 통해 외부로 빠져나가는 파일에 대한 실질적인 정보습득과 복원을 수행한다. 이러한 과정에서 유출된 문서의 최종 목적지 추적 및 파일 복원을 통한 해당 유출자에 대한 징계 근거자료로 제출될 수 있고다. 또한 공격패턴 분석을 통해 공격 시그니처를 정립하여 추후 공격에 대비 할 수 있다.

참고 문헌

[1] 최현철, “디지털 포렌식 심사절차에 관한 연구”, 2009. 군산대학교 석사학위논문

[2] 이글루스, “Pacp File Format”, http://mindori.egloos.com/2390488 (2016.03.03)

[3] 패킷인사이드, “PCAP 파일”, http://www.packetinside.com/ (2016.03.07)

[4] 위키백과, “컴퓨터 포렌식”, https://ko.wikipe%ED.orq/Wiki/%EC%BB%B4%ED%93%A8%ED%84%

B0\_%ED%8F%AC%EB%A0%8C%EC%8B%9D (2016.03.10)

[5] 전성훈, “첨단 바이오기술 이메일·USB로 줄줄 유출”, 연합뉴스, 2010.11.16 입력, http://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=102&oid=001&aid=0004767937

(2016.03.10)

[6] 노시영, "네트워크 포렌식을 위한 IDS 관제 방식 연구", 2010., 군산대학교 대학원